

Stochastic analysis in a tubular neighborhood or Onsager-Machlup functions revisited

(On the occasion of 70th birthday of Y.Takahashi)

Keisuke Hara* Yoichiro Takahashi

1 A personal introduction by K.H.

Inspired by K. Itô's work on Feynman path integrals and discussions on the effect of curvature among physicists, Y. Takahashi started studying the problem of the most probable path (or the Onsager-Machlup function) in late 1970's and completely determined the function in his joint work with S. Watanabe [6] (1980).

Soon after their work, T. Fujita and S. Kotani [1] (1982) gave the general result. While they used a singular perturbation method for P.D.E., Takahashi-Watanabe's proof is almost probabilistic except the only one analytic part, which treats asymptotic behaviour of a certain Wiener functional. Though this part gives a key observation, Takahashi removed it in a joint work with me to get a purely probabilistic proof and to extend the result to pinned diffusion processes [4] (1996).

On the other hand, he had been studying the Wiener functional in his way, and wrote a *draft*. In Takahashi-Watanabe [6], we can see how the scalar curvature term appears through the functional, while we cannot see the effect in [4] because the proof is too simple. Actually it appears from a kind of ergodic theorem due to the rapid rotation of the spherical motion as the tube is shrinking. His project was to study the effect with pure stochastic analysis.

Unfortunately, the draft had been missing for long years because he gave up the work to move to the other projects and also because "yet another" proof became somehow needless after the third proof.

However, recently I found the draft (written by himself with a primitive word processing software) in my library. Though it is hard to read because of his confusing writing, missing many parts, and the oxidization, I also found it unexpectedly interesting and tricky. Then, I tried to restore it as much as possible, reassembled, and I submit it as our *joint work* on the occasion of his 70th birthday after 20 years since he handed me the draft.

*Email: hara.keisuke@gmail.com

2 Introduction and the main result

Let $x(t)$ be a non-degenerate smooth diffusion process on a d -dimensional Riemannian manifold M and consider the asymptotics of the probability of the trajectory being confined in a small tubular neighbourhood of a prescribed smooth curve γ on M . The second author and S. Watanabe [6] proved that the following asymptotics is true. The functional L is the Onsager-Machlup function, which is similar to the Lagrangian. The point is the effect of the curvature to the “most probable path”.

Theorem 1

$$\lim_{\delta \rightarrow 0} \frac{P_{\gamma(0)} [\max_{0 \leq t \leq T} d(x(t), \gamma(t)) \leq \delta]}{P_0 [\max_{0 \leq t \leq T} |B(t)| \leq \delta]} = \exp\{-S(\gamma)\},$$

where $B(t)$ is the standard Brownian motion in the d -dimensional Euclidean space, $d(\cdot, \cdot)$ is the Riemannian distance on M ,

$$S(\gamma) = S_T(\gamma) = \int_0^T L(\gamma(t), \dot{\gamma}(t)) dt,$$

$$L(x, v) = \frac{1}{2} |f - v|_x^2 + \frac{1}{2} \operatorname{div} f(x) + \frac{1}{12} R(x),$$

and $R(x)$ is the Scalar curvature at x .

The theorem above was proved firstly by Takahashi-Watanabe [6]. In their almost probabilistic proof, they studied asymptotic behaviour of a certain Wiener functional to produce a key geometric quantity as an ergodic effect. Though this analysis is crucial, it is the only one “non-probabilistic” part in their proof. Soon after their work, Fujita-Kotani [1] gave an analytic proof to the theorem by a singular perturbation method for P.D.E.

Following these two proofs, the authors [4] gave another proof, which is simple and purely probabilistic, and extended the result to pinned diffusion processes. However, in return for the simplicity, the new proof lost a view to the ergodic effect. The aim of this paper is to give yet another (purely) probabilistic proof by recovering such precise study on the Wiener functional.

To highlight the difference between the old proofs and our new proof, here we recall the argument used in [6] and sketch our new idea.

The first key point is the existence of the Lagrangian. It follows from the two facts. The first fact is that for any given curve γ the law of the distance process $d(x(t), \gamma(t))$ is absolutely continuous to the law of a Bessel process with index $d = \dim M$ if they are restricted to a sufficiently small neighbourhood of the curve. We can find the law by adjusting the drift term expressed in normal coordinates along the curve γ , which turns out to be the difference of the Coriolis drift and the Besselization drift. Thus we find a diffusion process whose radial part admits the Bessel process and take it as the reference process to compute the Radon-Nykodim density given by Girsanov formula.

The latter fact is that the spherical motion of the reference process is governed by Levy's stochastic areas, which are orthogonal to the martingales governing the radial motion. Hence we can apply Kunita-Watanabe theorem to factorize the density into the Lagrangian and the asymptotically vanishing term.

The second key point is the derivation of the form of the Lagrangian $L(x, v)$, which consists of three terms as its stated above. The first term $\frac{1}{2}|v - f|_x^2$ comes out from the quadratic variation term in the Girsanov exponent. The other terms are obtained as the correction due to the transformation of the martingale term in the exponent into the stochastic linear integral. We apply the stochastic Stokes theorem to estimate the line integral.

The third term is the mean curvature term, which is hardest to obtain. Though it appears similarly as the difference of the quantities, it can be obtained more indirectly as an ergodic phenomenon. In Takahashi-Watanabe [6], this part (and only this part) was proved by appealing to an analytical method, i.e., the singular perturbation technique, using which Fujita-Kotani [1] gave another whole proof. After these two proofs, the authors gave a purely probabilistic method by "smooth Besselization technique" [4], which was introduced in [3] by the first author. However, it lost the concrete study of the ergodic effect in return for simplicity.

In this paper we prove this part by stochastic analysis. Precisely to say, the word "ergodic" stated above is not one in the usual sense but it is asymptotic. We can find a process that may be called asymptotically tangent to the original process as the tube shrinks, which is nothing but a Brownian motion. To justify this argument, we need to prepare a *strong* comparison theorem between diffusion processes with different diffusion constants. The comparison theorem guarantees that an asymptotic ergodic theorem is inherited by the original diffusion process from the asymptotically tangent process (i.e., a Brownian motion).

The trick for the strong comparison is to adjust the Brownian motions in the stochastic differential equations such that the distance between the two processes is kept small enough. We prepare a key lemma of linear algebra for this task.

3 Coriolis drift and Besselization drift

First we prepare necessary geometric quantities and introduce two important vectors: Coriolis drift and Besselization drift. Though we use properties of normal coordinates and basic asymptotics of the geometric tensors in the coordinates, we omit the proofs. They can be found in textbooks like [7] or [2].

The equation associated with the diffusion process $x(t)$ is, by the assumption, a non-degenerate second order parabolic differential equation on M as follows:

$$\frac{\partial u}{\partial t} = \frac{1}{2} \Delta_g u + fu,$$

where Δ_g is the Laplace-Beltrami operator for Riemannian metric g and f is a

vector field on M . In local coordinates they are expressed as

$$\Delta_g u(x) = \frac{1}{\sqrt{g(x)}} \sum_{i,j=1}^d \frac{\partial}{\partial x^j} \left(\sqrt{g(x)} g^{ij}(x) \frac{\partial u}{\partial x^i}(x) \right), \quad f u(x) = \sum_{i=1}^d f^i(x) \frac{\partial u}{\partial x^i}(x).$$

Here $(g^{ij}(x))$ is the diffusion coefficient in the coordinate, and its inverse matrix $(g_{ij}(x))$ gives the metric $ds^2 = \sum g_{ij}(x) dx^i dx^j$. We also used the conventional notation $g(x) = \det(g_{ij}(x))$. The norm $|\cdot|_x$ is the Riemannian norm on the tangent space $T_x M$ at the point x , i.e.,

$$|v|_x^2 = \sum_{i,j=1}^d g_{ij}(x) v^i v^j \quad \text{for} \quad v = \sum_{i=1}^d v^i \frac{\partial}{\partial x^i}.$$

The divergence of a vector field $f = \sum f^i(\partial/\partial x^i)$ is denoted by $\operatorname{div} f$, i.e.,

$$\operatorname{div} f(x) = \frac{1}{\sqrt{g(x)}} \sum_{i=1}^d \frac{\partial}{\partial x^i} \left(\sqrt{g(x)} f^i(x) \right).$$

We rewrite the Laplace-Beltrami operator as follows:

$$\frac{1}{2} \Delta_g u(x) = \sum_{i,j=1}^d \frac{1}{2} g^{ij}(x) \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} u(x) + \sum_{i=1}^d a^i(x) \frac{\partial}{\partial x^i} u(x)$$

where $a(x) = (a^i(x))$ is defined by

$$a^i(x) = \sum_{j=1}^d \frac{1}{\sqrt{g(x)}} \frac{\partial}{\partial x^j} \left(\sqrt{g(x)} g^{ij}(x) \right).$$

This drift term expressed by $a(x)$ is called Coriolis drift.

Let $\gamma(t)$ ($0 \leq t \leq T$) be a smooth curve on the manifold M . In the following, we take the lifting along the curve γ , i.e., normal coordinates along γ or Fermi coordinates (See [2]).

To write down our stochastic differential equation (S.D.E.), we prepare the square root $\sigma(t, x)$ of the matrix $(g^{ij}(t, x))$:

$$\sigma(t, x) = (\sigma^{ij}(t, x))_{i,j=1,\dots,d} = (\sigma_\alpha^i(t, x))_{i,\alpha=1,\dots,d} = (g^{ij}(t, x))^{1/2}.$$

Since $\sum_j g^{ij}(x) x^j = x^i$ holds in normal coordinates, we have $\sum_\alpha \sigma_\alpha^i(x) x^\alpha = x^i$. Note that we have also the same identities in Fermi coordinates, i.e., $\sum_j g^{ij}(t, x) x^j = x^i$ and $\sum_\alpha \sigma_\alpha^i(t, x) x^\alpha = x^i$.

Lemma 1 *There exists a local vector field $c(t, x)$ defined in normal coordinates along γ such that the solution $Y(t)$ to the S.D.E.*

$$dY(t) = \sigma(t, Y(t)) dB(t) + c(t, Y(t)) dt \quad (1)$$

admits the Bessel process of index d as its radial motion. Moreover, if, in addition, one requires the symmetric condition $c^i(t, x)x^j = c^j(t, x)x^i$ ($i, j = 1, \dots, d$), then it is uniquely determined and is given as

$$c^i(t, x) = \frac{x^i}{2|x|^2} \sum_{j=1}^d (1 - g^{jj}(t, x)).$$

Remark. The choice of this Besselization drift is a key point. The drift $c(t, x)$ has the coordinate symmetry, but it has also the singularity at $x = 0$. On the other hand, in [4] (originated in [3]) we broke the symmetry to choose another Besselization drift, which is totally smooth without any singularity.

Proof. Apply Itô formula to $\sum_{j=1}^d Y^j(t)^2$ and compare it with S.D.E. of the Bessel process. See [6] for details.

We have the asymptotic expansion of our geometric tensors $\sigma(t, x)$, $a(t, x)$, and $c(t, x)$ as $x \rightarrow 0$ in normal coordinates along the curve as follows.

Lemma 2

$$\begin{aligned} \sigma^{ij}(t, x) &= \delta^{ij} + \frac{1}{6} R_{ikjl}(t, 0) x^k x^l + O(|x|^3), \\ \sum_{j=1}^d \frac{\partial}{\partial x^j} a^j(t, x) &= -\frac{1}{3} R(t, 0) + O(|x|^2), \\ \sum_{j=1}^d \frac{\partial}{\partial x^j} c^j(t, x) &= \frac{d}{6} \sum_{i,j=1}^d R_{ij}(t, 0) u^i u^j + O(|x|), \end{aligned}$$

where δ^{ij} is the Kronecker delta, R_{ijkl} is the Riemannian curvature, R_{ij} is the Ricci curvature, R is the scalar curvature, and $u^i = x^i/|x|$.

4 Girsanov formula

In the same way as the preceding papers [6] [1] [4], we write the S.D.E. of the process $X(t)$ obtained by lifting the original process $x(t)$ as follows.

$$dX^i(t) = \sum_{\alpha=1}^d \sigma_{\alpha}^i(t, X(t)) dB^{\alpha}(t) + \tilde{b}^i(t, X(t)) dt, \quad (2)$$

where the drift term \tilde{b} is

$$\tilde{b}^i(t, x) = a^i(t, x) + f^i(t, x) - \dot{\gamma}^i(t)$$

in normal coordinates along $\gamma(t)$. To use it later, we set

$$b^i(t, x) = f^i(t, x) - \dot{\gamma}^i(t).$$

Let us compare $X(t)$ with $Y(t)$ whose radial part is a Bessel process. By Girsanov formula, we obtain the following (For the proof, see [6] or [4]).

Proposition 1 As $\delta \rightarrow 0$,

$$\begin{aligned} P_{\gamma(0)} \left[\max_{0 < t < T} d(x(t), \gamma(t)) < \delta \right] &= P_0 \left[\max_{0 < t < T} d(X(t), 0) < \delta \right] \\ &= \exp\{-S_T(\gamma) + O(\delta)\} E_0 \left[\exp \left(\int_{Y[0,T]} \alpha + \int_0^T \beta(t) dt \right); |Y|_T < \delta \right], \end{aligned}$$

where $|Y|_T = \max_{[0,T]} |Y(t)|$, the integral over $Y[0, T]$ is the stochastic line integral over the space-time curve $(t, \gamma(t))$, $0 \leq t \leq T$, the integrand α is the space-time 1-form (which is degenerated in t -direction) given by

$$\alpha = \sum_{i=1}^d \alpha_i dx^i \quad \text{with} \quad \alpha_i = \sum_{j=1}^d (a^j(t, x) + b^j(t, x) - c^j(t, x)) g_{ij}(t, x),$$

the integrand $\beta(t)$ stands for the function

$$\beta(t) = \frac{d}{12} \sum_{i,j=1}^d R_{ij}(\gamma(t)) \left(\frac{Y^i(t)}{|Y(t)|} \frac{Y^j(t)}{|Y(t)|} - \frac{\delta^{ij}}{d} \right).$$

5 Stochastic Stokes theorem

Now let us apply the stochastic version of the Stokes theorem ([5]) to get the estimate of the stochastic line integral of the 1-form α over the space-time curve $Y[0, t]$. Let Σ be the (random) surface $\{sY(t)\}_{0 \leq s \leq 1, 0 \leq t \leq T}$. Then its boundary $\partial\Sigma$ consists of four curves: the trajectory $Y[0, T]$ (corresponding to $s = 1$), the segment from $(T, Y(T))$ to $(T, 0)$ (corr. to $t = T$), the segment from $(T, 0)$ to $(0, 0)$ (corr. to $t = 0$), and the degenerated segment consisting of one point $(0, 0)$ (corr. to $s = 0$).

Note that the line integrals over those segments are of order $O(\delta)$ on the set where $|Y|_T < \delta$ because $d(sY(t)) = Y(t)ds$ on the segments where t is a constant. Then, stochastic Stokes theorem says that

$$\int_{\Sigma} d\alpha = \int_{\partial\Sigma} \alpha = \int_{Y[0,T]} \alpha + O(\delta T).$$

Therefore, the estimate of the line integral is deduced to the study of the area, which is expressed concretely as follows.

Lemma 3 Denote the stochastic areas (Lévy's area) by

$$A^{ij}(t) = \int_0^t (Y^i \circ dY^j - Y^j \circ dY^i)$$

and set

$$\alpha_{ij}(t, x) = \frac{1}{4} \int_0^1 \{\partial_i \alpha_j(t, sx) - \partial_j \alpha_i(t, sx)\} ds.$$

Then,

$$\int_{\Sigma} d\alpha = \sum_{i,j=1}^d \int_0^T \alpha_{ij}(t, Y(t)) \circ dA^{ij}(t).$$

Proof. Obvious from the definitions. (Notice that we omitted the factor $(1/2)$ from A^{ij} , which is set into α_{ij} .)

The following lemma is crucial, because we need the orthogonality to factorize the density into the Lagrangian term and the asymptotically vanishing term by Kunita-Watanabe theorem.

Lemma 4 *The stochastic areas $A^{ij}(t)$ are martingales which are orthogonal to the radial motion $|Y(t)|$. In other words, the quadratic variation process vanishes:*

$$\langle A^{ij}, |Y| \rangle(t) = 0.$$

Proof. Show $(Y^i \circ dY^j - Y^j \circ dY^i) \sum_{k=1}^d Y^k dY^k = 0$ by stochastic calculus. See [6] or [4].

Therefore we proceed to the next step as follows:

Proposition 2 *For $p > 0$, set*

$$M_p(t) = p \int_0^t \alpha_{ij}(s, Y(s)) dA^{ij}(s) - \frac{p^2}{2} \int_0^t \alpha_{ij} \alpha_{kl} d\langle A^{ij}, A^{kl} \rangle(s), \quad M(t) = M_1(t),$$

and

$$L(t) = \frac{d}{12} \int_0^t R_{ij}(\gamma(s)) \left\{ U^i(s) U^j(s) - \frac{\delta^{ij}}{d} \right\} ds,$$

where $U(t)$ is the spherical part of $Y(t)$, i.e., $U^i(t) = Y^i(t)/|Y(t)|$. Then,

$$\begin{aligned} & P \left[\max_{0 \leq t \leq T} d(x(t), \gamma(t)) < \delta \right] \\ &= \exp \{ -S_T(\gamma) + O(\delta T) \} E [\exp \{ M(t) + L(t) \}; |Y|_T < \delta]. \end{aligned}$$

Proof. Note that on the set $|Y|_T < \delta$,

$$\int_0^T \alpha_{ij} \alpha_{kl} d\langle A^{ij}, A^{kl} \rangle(t) = O(\delta^2 T),$$

because $(Y^i \circ Y^j - Y^j \circ Y^i)(Y^k \circ Y^l - Y^l \circ Y^k) = O(\delta^2) dt$. And also,

$$\langle \alpha_{ij}, A^{ij} \rangle(T) = O(\delta T),$$

because $d\alpha_{ij}(Y^i \circ Y^j - Y^j \circ Y^i) = O(\delta) dt$. Hence,

$$\int_0^T \alpha_{ij} \circ dA^{ij} = M(T) + O(\delta T) \quad \text{on } |Y|_T < \delta.$$

Consequently the assertion follows from Proposition 1.

Remark. If (M, g) is an Einstein space, we have the relation $R_{ij} = (1/d)g_{ij}$ and so $L(t) = 0$. Therefore, applying Kunita-Watanabe theorem to the exponential martingale part, we already have proved Theorem 1 in this case.

6 Pathwise adjustment

The following sections are the new part where we study the Wiener functional $L(t)$ with stochastic analysis.

Let us recall that the process $Y(t)$ is governed by the S.D.E.:

$$dY(t) = \sigma(t, Y(t))dB(t) + \frac{Y(t)}{2|Y(t)|^2} \left(d - \sum_{j=1}^d g^{jj}(t, Y(t)) \right) dt.$$

Since $\sigma(t, Y(t)) = I + O(|Y(t)|^2)$, one may expect that the process $Y(t)$ converges as $\delta \rightarrow 0$ under $P[\cdot | |Y(t)|_T < \delta]$ to a d -dimensional Brownian motion, possibly with a drift, in some sense and that we can estimate its convergence rate. However, usual comparison theorems for S.D.E.s give very crude information when the diffusion coefficients vary in contrast with comparison theorems for partial differential equations. Moreover what we want to deal with is its spherical part $U(t) = Y(t)/|Y(t)|$. We solve these problems as follows.

By Itô formula, the spherical part $U(t)$ is governed by the S.D.E.:

$$\begin{aligned} dU(t) &= \frac{1}{|Y(t)|} (I - U(t) \otimes U(t)) \sigma(t, Y(t)) dB(t) \\ &\quad - \frac{1}{2|Y(t)|^2} \left(\sum_{j=1}^d g^{jj}(t, Y(t)) - 1 \right) U(t) dt, \end{aligned} \quad (3)$$

where we denote by $u \otimes v$ the matrix $(u^i v^j)_{1 \leq i, j \leq d}$.

On the other hand, the spherical part $\tilde{U}(t)$ of a d -dimensional standard Brownian motion $\tilde{B}(t)$ is, of course, governed by the S.D.E.:

$$d\tilde{U}(t) = \frac{1}{|\tilde{B}(t)|} \left(I - \tilde{U}(t) \otimes \tilde{U}(t) \right) d\tilde{B}(t) - \frac{d-1}{2|\tilde{B}(t)|^2} \tilde{U}(t) dt. \quad (4)$$

The idea to obtain a good comparison estimate between (3) and (4) is to adjust the Brownian motions so that the radial part $|\tilde{B}(t)|$ of $\tilde{B}(t)$ is *pathwise* equal to the radial part $|Y(t)|$ of $Y(t)$. With this trick, we can manage the fluctuation of the spherical part enlarged by the fluctuation of radial parts. For this sake we prepare the following lemma.

Lemma 5 *Let $n = 1 + d(d-1)/2$. Then there are linear maps J^i ($i = 1, \dots, d$) from \mathbf{R}^d to \mathbf{R}^n with the following properties:*

(a) Each J^i is an isometry of \mathbf{R}^d into \mathbf{R}^n . In other words, we have

$$\langle J^i u, J^i v \rangle_n = \langle u, v \rangle_d$$

for any u, v in \mathbf{R}^d , where $\langle \cdot, \cdot \rangle_n$ and $\langle \cdot, \cdot \rangle_d$ denote the inner products in \mathbf{R}^n and \mathbf{R}^d respectively.

(b) For each unit vector $u \in \mathbf{R}^d$, $\{J^i u\} (1 \leq i \leq d)$ are orthonormal, i.e.,

$$\langle J^i u, J^j u \rangle_n = \delta^{ij}.$$

(c) There is a unit vector $e_0 \in \mathbf{R}^n$ such that

$$\sum_{i=1}^d u^i J^i u = |u|_d^2 e_0$$

for every $u = (u^i)_{1 \leq i \leq d}$ in \mathbf{R}^d .

Proof. Let $e_\alpha (1 \leq \alpha \leq d)$ be the canonical basis in \mathbf{R}^d and $e_0, e_{\alpha\beta} (\alpha < \beta)$ be an orthonormal basis in \mathbf{R}^n . (Thus, $n = 1 + d(d-1)/2$.) Set

$$J^i e_\alpha = \begin{cases} e_{i\alpha} & \text{if } i < \alpha, \\ e_0 & \text{if } i = \alpha, \\ -e_{\alpha i} & \text{if } i > \alpha. \end{cases}$$

Then a direct, elementary computation shows that

$$\langle J^i e_\alpha, J^j e_\beta \rangle = \delta_{ij} \delta_{\alpha\beta} - \delta_{i\beta} \delta_{\alpha j} + \delta_{\alpha i} \delta_{j\beta}.$$

Now from the equation above one obtains (a) by putting $i = j$ and (b) by putting $\alpha = \beta$. Finally,

$$\begin{aligned} \sum_{i=1}^d u^i J^i u &= \sum_{i=1}^d \sum_{\alpha=1}^d u^i u^\alpha J^i e_\alpha \\ &= \sum_{i < \alpha} u^i u^\alpha e_{i\alpha} + \sum_i (u^i)^2 e_0 + \sum_{i > \alpha} -u^i u^\alpha e_{\alpha i} = |u|^2 e_0. \end{aligned}$$

Hence we obtain (c).

Now let us take a $(1 + d(d-1)/2)$ -dimensional Brownian motion W and consider the following S.D.E.s:

$$\begin{aligned} dY(t) &= \sigma(t, Y(t)) \langle JU(t), dW(t) \rangle - \frac{1}{2|Y(t)|^2} \sum_{j=1}^d \{g^{jj}(t, Y(t)) - 1\} Y(t) dt, \\ d\tilde{Y}(t) &= \langle J\tilde{U}(t), dW(t) \rangle, \end{aligned}$$

where $U(t) = Y(t)/|Y(t)|$, $\tilde{U}(t) = \tilde{Y}(t)/|\tilde{Y}(t)|$, and $\langle Ju, v \rangle$ stands for the d -dimensional vector whose components are $\langle J^i u, v \rangle_n (1 \leq i \leq d)$. Solve these

S.D.E.s above and then we can define the desired Brownian motions $B(t)$ and $\tilde{B}(t)$ by

$$dB(t) = \langle JU(t), dW(t) \rangle \quad \text{and} \quad d\tilde{B}(t) = \langle J\tilde{U}(t), dW(t) \rangle.$$

Then it is obvious that

$$\langle U(t), dB(t) \rangle = \langle \tilde{U}(t), d\tilde{B}(t) \rangle = dW_0,$$

where $W_0(t)$ is the 1-dimensional Brownian motion which is defined by

$$W_0(t) = \int_0^t \langle e_0, dW(s) \rangle.$$

With these Brownian motions we obtain the desired S.D.E.s (3) and (4) for which, by the property (c) of Lemma 5, the radial parts coincide with each other if the initial data do so. In fact,

$$\begin{aligned} d(|Y(t)|^2) &= 2\langle Y(t), dY(t) \rangle + \langle dY(t), dY(t) \rangle \\ &= 2\langle Y(t), \sigma(t, Y(t))dB(t) \rangle \\ &\quad - 2\langle Y(t), \frac{1}{2|Y(t)|^2} \sum_{j=1}^d (g^{jj} - 1)Y(t)dt \rangle + \sum_{j=1}^d g^{jj} dt \\ &= 2\langle Y(t), dB(t) \rangle + d \cdot dt \\ &= 2|Y(t)|\langle U(t), dB(t) \rangle + d \cdot dt = 2|Y(t)|dW_0(t) + d \cdot dt. \end{aligned}$$

On the other hand,

$$\begin{aligned} d(|\tilde{Y}(t)|^2) &= 2\langle \tilde{Y}(t), d\tilde{B}(t) \rangle + d \cdot dt = 2|\tilde{Y}(t)|\langle \tilde{U}(t), d\tilde{B}(t) \rangle + d \cdot dt \\ &= 2|\tilde{Y}(t)|dW_0(t) + d \cdot dt. \end{aligned}$$

Hence, $|Y(t)| = |\tilde{Y}(t)|$ if $|Y(0)| = |\tilde{Y}(0)|$.

7 The inner product of the spherical parts

Next let us deduce the S.D.E. for the inner product of $U(t)$ and $\tilde{U}(t)$.

Lemma 6 *The inner product $\langle U(t), \tilde{U}(t) \rangle$ satisfies the following S.D.E.:*

$$d\langle U, \tilde{U} \rangle = R(t)H(t)dW_1(t) - \frac{1}{2}R(t)^2G(t)\langle U, \tilde{U} \rangle dt,$$

where $W_1(t)$ is a 1-dimensional Brownian motion that is independent of $W_0(t)$,

$$H(t) = H(t, R(t), U(t), \tilde{U}(t)) = \frac{1}{R(t)^2} \left| (\sigma(t, R(t)U(t)) - I) (U(t) - \tilde{U}(t)) \right|,$$

$$G(t) = G(t, R(t), U(t)) = \frac{1}{R(t)^4} \text{tr} \left((\sigma(t, R(t)U(t)) - I)^2 \right).$$

Proof. For a while let us use the notations,

$$\begin{aligned}\tau &= \tau(t, R(t), U(t)) = \sigma(t, R(t)U(t)) - U(t) \otimes U(t), \\ \tilde{\tau} &= \tilde{\tau}(\tilde{U}(t)) = I - \tilde{U}(t) \otimes \tilde{U}(t).\end{aligned}$$

Then the S.D.E.s governing $U(t)$ and $\tilde{U}(t)$ can be written as

$$dU = \frac{\tau}{R}dB - \frac{\text{tr}(\tau^2)}{2R^2}Udt \quad \text{and} \quad d\tilde{U} = \frac{\tilde{\tau}}{R}d\tilde{B} - \frac{\text{tr}(\tilde{\tau}^2)}{2R^2}\tilde{U}dt.$$

Recall that the Brownian motions $B(t)$ and $\tilde{B}(t)$ are adjusted so that

$$dB = \omega(t)d\tilde{B}(t) \quad \text{and} \quad U(t) = \omega(t)\tilde{U}(t),$$

where $\omega(t) = (\omega^{ij}(t))$ is the unitary matrix which is given by

$$\omega(t) = U(t) \otimes \tilde{U}(t) - \tilde{U}(t) \otimes U(t) + \langle U(t), \tilde{U}(t) \rangle I.$$

Consequently,

$$\begin{aligned}d\langle U, \tilde{U} \rangle &= \langle \tilde{U}, \frac{\tau}{R}dB \rangle - \frac{\text{tr}(\tau^2)}{2R^2}\langle \tilde{U}, U \rangle dt + \langle U, \frac{\tilde{\tau}}{R}d\tilde{B} \rangle \\ &\quad - \frac{\text{tr}(\tilde{\tau}^2)}{2R^2}\langle U, \tilde{U} \rangle dt + \langle \frac{\tau}{R}dB, \frac{\tilde{\tau}}{R}d\tilde{B} \rangle \\ &= \frac{1}{R}\langle {}^t\omega {}^t\tau \tilde{U} + \tilde{\tau}U, d\tilde{B} \rangle - \frac{1}{2}R^2Gdt,\end{aligned}$$

where we put

$$G = \frac{1}{R^2} \left\{ (\text{tr}(\tau^2) + \text{tr}(\tilde{\tau}^2)) \langle U, \tilde{U} \rangle - 2\text{tr}(\tilde{\tau}\tau\omega) \right\}. \quad (5)$$

Soon we shall show that this quantity G coincides with the function $G(t)$ in the statement of the lemma.

Now let us compute the diffusion coefficients:

$$\begin{aligned}{}^t\omega {}^t\tau \tilde{U} + \tilde{\tau}U &= {}^t\omega(\sigma - I)\tilde{U} + {}^t\omega\tilde{U} - {}^t\omega\langle U, \tilde{U} \rangle U + U - \langle U, \tilde{U} \rangle \tilde{U} \\ &= {}^t\omega(\sigma - I)\tilde{U} + {}^t\omega\tilde{U} + \omega\tilde{U} - 2\langle U, \tilde{U} \rangle \tilde{U} = {}^t\omega(\sigma - I)\tilde{U}\end{aligned}$$

because ${}^t\omega + \omega = 2\langle U, \tilde{U} \rangle I$. Hence we can define a 1-dimensional Brownian motion $W_1(t)$ through the relation

$$\langle {}^t\omega {}^t\tau \tilde{U} + \tau\tilde{U}, d\tilde{B} \rangle = \langle (\sigma - I)\tilde{U}, dB \rangle = R(t)^2 H dW_1,$$

where we put

$$H = \frac{1}{R(t)^2} |(\sigma - I)\tilde{U}(t)| = \frac{1}{R(t)^2} |(\sigma - I)(U(t) - \tilde{U}(t))|.$$

Note that $dW_0 dW_0 = 0$. In fact, by virtue of $(\sigma - I)U = 0$, we have

$$\langle (\sigma - I)\tilde{U}, dB \rangle \langle U, dB \rangle = \langle (\sigma - I)\tilde{U}, U \rangle dt = \langle \tilde{U}, (\sigma - I)U \rangle dt = 0.$$

Consequently, the Brownian motions W_0 and W_1 are independent.

Lastly we check G in (5) coincides with $G(t)$ in the statement of the lemma. Note that

$$\begin{aligned} \text{tr}(\tilde{\tau}\tau\omega) &= \text{tr} \left\{ \tilde{\tau}\tau \left(U \otimes \tilde{U} - \tilde{U} \otimes U + \langle U, \tilde{U} \rangle I \right) \right\} \\ &= \langle \tilde{\tau}\tau U, \tilde{U} \rangle - \langle \tilde{\tau}\tau \tilde{U}, U \rangle + \langle U, \tilde{U} \rangle \text{tr}(\tilde{\tau}\tau) \\ &= \left(\langle \tau U, \tilde{U} \rangle - \langle \tilde{U} \otimes \tilde{U} \tau U, \tilde{U} \rangle \right) - \left(\langle \tau \tilde{U}, U \rangle - \langle \tilde{U} \otimes \tilde{U} \tau U, U \rangle \right) \\ &\quad + \langle U, \tilde{U} \rangle \text{tr} \left(\tau(I - \tilde{U} \otimes \tilde{U}) \right) \\ &= 0 - \left(\langle (\sigma - U \otimes U) \tilde{U}, U \rangle - \langle \tilde{U}, \tau \tilde{U} \rangle \langle \tilde{U}, U \rangle \right) + \langle U, \tilde{U} \rangle \left(\text{tr}(\tau) - \langle \tau \tilde{U}, \tilde{U} \rangle \right) \\ &= \langle U, \tilde{U} \rangle \text{tr}(\tau) = \langle U, \tilde{U} \rangle (\text{tr}(\sigma) - 1). \end{aligned}$$

Hence we obtain

$$\begin{aligned} R^2 G &= \{ (\text{tr}(\sigma^2) - 1) + (d - 1) \} \langle U, \tilde{U} \rangle - 2 \langle U, \tilde{U} \rangle (\text{tr}(\sigma) - 1) \\ &= \{ \text{tr}(\sigma^2) - 2 \text{tr}(\sigma) + d \} \langle U, \tilde{U} \rangle = \text{tr}((\sigma - I)^2) \langle U, \tilde{U} \rangle \end{aligned}$$

as is desired. The proof is completed.

8 The asymptotic analysis of the spherical parts

Now we are ready to prove that the spherical motion $U(t)$ is asymptotic to $\tilde{U}(t)$ “on the exponent”.

Proposition 3 *For $U(t)$ and $\tilde{U}(t)$ defined above, the conditional expectation*

$$E \left[\exp \left\{ \frac{c}{\sqrt{\delta}} |U(t) - \tilde{U}(t)| \right\} \middle| \max_{[0, T]} R(t) < \delta \right]$$

remains bounded as δ tends to 0 for every constant c .

Proof. Keeping in mind that $|U - \tilde{U}|^2 = 2(1 - \langle U, \tilde{U} \rangle)$, let us consider the process

$$\Delta(t) = (1 - \langle U(t), \tilde{U}(t) \rangle) \exp \left\{ \int_0^t \frac{1}{2} R(s)^2 G(s) ds \right\}. \quad (6)$$

It is immediate to see that it satisfies the stochastic differential equation

$$\begin{aligned} d\Delta(t) &= \left\{ -d\langle U, \tilde{U} \rangle + (1 - \langle U, \tilde{U} \rangle) \frac{1}{2} R^2 G dt \right\} \exp \int_0^t \frac{1}{2} R^2 G ds \\ &= -R(t) H(t) \left\{ \exp \int_0^t \frac{1}{2} R^2 G ds \right\} dW_1(t) + d \left(\exp \int_0^t \frac{1}{2} R^2 G ds \right). \end{aligned}$$

Now let us take a 1-dimensional Brownian motions W_2 , which is independent of the Bessel process $R(t)$, such that

$$-\int_0^t R(s)H(s) \left\{ \exp \int_0^s \frac{1}{2} R(r)^2 G(r) dr \right\} dW_1(s) = W_2(\nu(t))$$

where for simplicity of the notation we set

$$\nu(t) = \int_0^t R(s)^2 H(s)^2 \left\{ \exp \int_0^s R(r)^2 G(r) dr \right\} ds.$$

Therefore we can rewrite $\Delta(t)$ as follows:

$$\Delta(t) = W_2(\nu(t)) + \left(\exp \int_0^t \frac{1}{2} R(s)^2 G(s) ds - 1 \right). \quad (7)$$

Notice that

$$\nu(t) \leq \int_0^t R(s)^2 G(s) \left\{ \exp \int_0^s R(r)^2 G(r) dr \right\} ds = \exp \int_0^t R(s)^2 G(s) ds - 1,$$

since we have

$$H(t)^2 = \frac{1}{R(t)^4} \left| (\sigma - I) \tilde{U}(t) \right|^2 \leq \frac{\text{tr}((\sigma - I)^2)}{R(t)^4} = G(t).$$

Also notice that $G(t) < M$ for a constant M because $\sigma(t, x) - I = O(|x|^2)$.

Instead of the conditional expectation in the statement, let us estimate the following conditional probability

$$P_\delta(\lambda) = P \left[\sup_{[0, T]} \frac{1}{\sqrt{\delta}} \left| U(t) - \tilde{U}(t) \right| > \lambda \mid \max_{[0, T]} R(t) < \delta \right]$$

for $\lambda > 0$. Since $|U - \tilde{U}|/\sqrt{\delta} > \lambda$ is equivalent to $2(1 - \langle U, \tilde{U} \rangle)/\delta > \lambda^2$, by the definition (6) of $\Delta(t)$ we have

$$\begin{aligned} P_\delta(\lambda) &= P \left[\sup_{[0, T]} \frac{2}{\delta} \Delta(t) e^{-\int_0^t (1/2) R^2(s) G(s) ds} > \lambda^2 \mid \max_{[0, T]} R(t) < \delta \right] \\ &\leq P \left[\sup_{[0, T]} \Delta(t) > \frac{\lambda^2 \delta}{2} \mid \max_{[0, T]} R(t) < \delta \right]. \end{aligned}$$

Substituting (7) for $\Delta(t)$,

$$\begin{aligned} P_\delta(\lambda) &\leq P \left[\sup_{[0, T]} \left\{ W_2(\nu(t)) + \left(e^{\int_0^t (1/2) R^2(s) G(s) ds} - 1 \right) \right\} > \frac{\delta \lambda^2}{2} \mid \max_{[0, T]} R(t) < \delta \right] \\ &\leq P \left[\sup_{[0, T]} W_2(\nu(t)) + \left(e^{(1/2) \delta^2 M T} - 1 \right) > \frac{\delta \lambda^2}{2} \mid \max_{[0, T]} R(t) < \delta \right] \end{aligned}$$

$$\begin{aligned}
&\leq P \left[\sup_{[0,T]} W_2 \left(e^{\int_0^t R(s)^2 G(s) ds} - 1 \right) > \frac{\delta \lambda^2}{2} - \left(e^{(1/2)\delta^2 MT} - 1 \right) \middle| \max_{[0,T]} R(t) < \delta \right] \\
&\leq P \left[\sup_{[0, \exp(\delta^2 MT) - 1]} W_2(t) > \frac{\delta \lambda^2}{2} - \left(e^{(1/2)\delta^2 MT} - 1 \right) \middle| \max_{[0,T]} R(t) < \delta \right] \\
&= 2P \left[W_2(e^{\delta^2 MT} - 1) > \frac{\delta \lambda^2}{2} - \left(e^{(1/2)\delta^2 MT} - 1 \right) \right].
\end{aligned}$$

We used the Andre reflection principle and the independence of W_2 and R in the last equality. Therefore we have the final estimate

$$P_\delta(\lambda) \leq 2P \left[W_2(1) > \frac{\delta \lambda^2}{2} \left(e^{\delta^2 MT} - 1 \right)^{-1/2} - \left(e^{(1/2)\delta^2 MT} - 1 \right) \left(e^{\delta^2 MT} - 1 \right)^{-1/2} \right]$$

by the scale invariance of the Brownian motion. Since we can choose a constant K independently on δ such that

$$P_\delta(\lambda) \leq 2P [W_2(1) > K\lambda^2]$$

for large enough λ , the conditional expectation

$$E \left[\exp \left\{ \frac{c}{\sqrt{\delta}} |U(t) - \tilde{U}(t)| \right\} \middle| \max_{[0,T]} R(t) < \delta \right]$$

is bounded in $\delta > 0$ for every constant c . The proof is completed.

9 The last step to the main theorem

Now we can finish the proof of the main theorem.

By Proposition 2 it suffices for the proof of Theorem 1 to show that

$$E [\exp\{M(t) + L(t)\} \mid |Y|_T < \delta] \rightarrow 1 \quad \text{as } \delta \rightarrow 0.$$

In the following, we will show that its order is $\exp O(\sqrt{\delta}T)$.

Recall the following notations:

$$M_p(t) = p \int_0^t \alpha_{ij}(s, Y(s)) dA^{ij}(s) - \frac{p^2}{2} \int_0^t \alpha_{ij} \alpha_{kl} d\langle A^{ij}, A^{kl} \rangle(s), \quad M(t) = M_1(t),$$

$$L(t) = \int_0^t \beta(s, U(s)) ds \quad \text{with} \quad U(t) = \frac{Y(t)}{|Y(t)|},$$

and

$$\beta(t, u) = \frac{d}{12} \sum_{i,j=1}^d R_{ij}(\gamma(t)) \left\{ u^i u^j - \frac{\delta^{ij}}{d} \right\}.$$

Set

$$\tilde{L}(t) = \int_0^t \beta(s, \tilde{U}(s)) ds$$

where $\tilde{U}(t) = \tilde{B}(t)/|B(t)|$ is the spherical Brownian motion introduced before. Since $\beta(t, u)$ is Lipschitz with respect to u uniformly in t on $[0, T]$, it follows from Proposition 3 that there is a constant C_1 such that

$$E \left[\exp \left\{ \frac{1}{\sqrt{\delta}} (L(t) - \tilde{L}(t)) \right\} \middle| |Y|_T < \delta \right] < C_1. \quad (8)$$

We have also that there exists a constant C_2 such that

$$E \left[\exp \left\{ \frac{1}{\sqrt{\delta}} \tilde{L}(T) \right\} \middle| |Y|_T < \delta \right] < C_2. \quad (9)$$

Now we apply Hölder inequality to our target expectation after a little modification. First of all, note that for any positive number p , we have

$$\begin{aligned} & E [\exp(M(t) + L(t)) \mid |Y|_T < \delta] \\ &= \exp(O(\delta T)) E \left[\exp \left\{ \frac{1}{p} M_p(t) + L(t) \right\} \middle| |Y|_T < \delta \right]. \end{aligned} \quad (10)$$

Now take small δ and set $p = (1 - 2\sqrt{\delta})^{-1}$. Then, $1/p + \sqrt{\delta} + \sqrt{\delta} = 1$ and so

$$\begin{aligned} & E \left[\exp \left\{ \frac{1}{p} M_p(t) + L(t) \right\} \middle| |Y|_T < \delta \right] \\ &\leq E [\exp M_p(t) \mid |Y|_T < \delta]^{1/p} \cdot E \left[\exp \left\{ \frac{1}{\sqrt{\delta}} (L(t) - \tilde{L}(t)) \right\} \middle| |Y|_T < \delta \right]^{\sqrt{\delta}} \\ &\quad \times E \left[\exp \left\{ \frac{1}{\sqrt{\delta}} \tilde{L}(t) \right\} \middle| |Y|_T < \delta \right]^{\sqrt{\delta}} \leq C_1^{\sqrt{\delta}} C_2^{\sqrt{\delta}} = \exp O(\sqrt{\delta} T). \end{aligned} \quad (11)$$

Here we used the fact that $\exp M_p(t)$ is a martingale under the conditional probability $P[\cdot \mid |Y|_T < \delta]$. From (10) and (11), we obtain the upper estimate

$$E [\exp \{M(t) + L(t)\} \mid |Y|_T < \delta] < \exp O(\sqrt{\delta} T).$$

The inverse inequality is immediately obtained by Jensen inequality. In fact,

$$E [\exp \{M(t) + L(t)\} \mid |Y|_T < \delta] \geq \exp E [M(t) + L(t) \mid |Y|_T < \delta].$$

It is obvious that

$$\exp E [M(t) \mid |Y|_T < \delta] = \exp O(\delta T).$$

On the other hand, it follows from (9) and (8) using Proposition 3 that

$$\begin{aligned} & \exp E [L(t) \mid |Y|_T < \delta] \\ &= \exp E [\tilde{L}(T) \mid |Y|_T < \delta] \exp E [L(t) - \tilde{L}(t) \mid |Y|_T < \delta] \\ &= \exp O(\sqrt{\delta} T) \exp O(\sqrt{\delta} T) = \exp O(\sqrt{\delta} T). \end{aligned}$$

Consequently,

$$E [\exp \{M(t) + L(t)\} \mid |Y|_T < \delta] \geq \exp O(\sqrt{\delta} T).$$

Hence the proof of the Theorem is completed.

References

- [1] T. Fujita and S. Kotani. “The Onsager-Machlup functions for diffusion processes” J.Math.Kyoto. Univ., **22-1** (1982), pp.115–130.
- [2] R. Gray. “Tube”, Addison-Wesley Publishing Company, (1990).
- [3] K. Hara. ”Wiener functionals associated with joint distributions of exit time and position from small geodesic balls”, The Annals of Probability, Vol.24, No.2, (1996), pp.825–837.
- [4] K. Hara and Y. Takahashi. “Lagrangian for pinned diffusion process” in “Itô’s Stochastic Calculus and Probability Theory”, Springer (1996), pp.117–128.
- [5] N. Ikeda and S. Watanabe. “Stochastic Differential Equations and Diffusion Processes” (2nd Ed.), North-Holland/Kodansha (1989).
- [6] Y. Takahashi and S. Watanabe. “The probability functionals (Onsager-Machlup functions) of diffusion processes”, Proc.L.M.S.Symp. on Stochastic Integrals at Durham, Lecture Note in Math, **851** (1980), pp.443–463.
- [7] T. J. Willmore. “Riemannian Geometry”, Oxford University Press (1993).

Keisuke HARA
Mynd, Inc.
Shirokanedai Bld. 3F
3-2-10 Shirokanedai, Minato-ku,
Tokyo, 108-0071 JAPAN
Email: hara.keisuke@gmail.com